Direct Estimation of the Derivative of Quadratic Mutual Information with Application in Sufficient Dimension Reduction

Voot Tangkaratt (Tokyo Tech), Hiroaki Sasaki (UTokyo), and Masashi Sugiyama (UTokyo)

Sufficient dimension reduction (SDR)

- Projection matrix W is often obtained by *gradient* ascent:

$$W \leftarrow W + \varepsilon \nabla J(W)$$

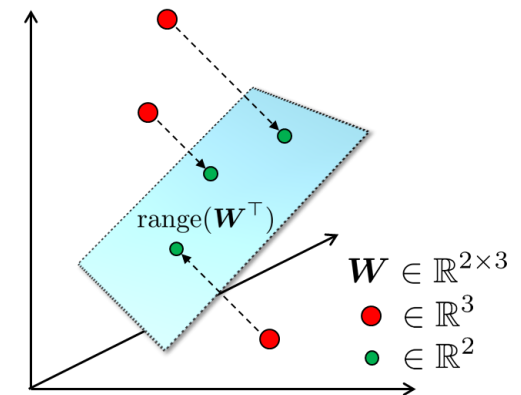 - $J(W)$ is some SDR criterion.
    - E.g., *Quadratic Mutual Information*

Existing method

- Calculates derivative of the estimated $\hat{J}(W)$: $\nabla \hat{J}(W)$

Propose method

- Directly estimates the derivative: $\widehat{\nabla J(W)}$

$$\text{range}(W^\top)$$

$W \in \mathbb{R}^{2 \times 3}$
$\bullet \in \mathbb{R}^3$
$\bullet \in \mathbb{R}^2$

Linear projection of 3 data points by matrix W.